



日本の研究者によるデータ公開の実践状況と認識

著者	池内 有為
発行年	2019
学位授与大学	筑波大学 (University of Tsukuba)
学位授与年度	2019
報告番号	12102乙第2937号
URL	http://hdl.handle.net/2241/00158191

氏 名	池内有為
学 位 の 種 類	博 士（図書館情報学）
学 位 記 番 号	博 乙 第 2937 号
学位授与年月日	令和元年9月25日
学位授与の要件	学位規則第4条第2項該当
審 査 研 究 科	図書館情報メディア研究科
学位論文題目	日本の研究者によるデータ公開の実践状況と認識

主 査	筑波大学	教授	文学修士	逸村裕
副 査	筑波大学	教授	博士（政治学）	溝上智恵子
副 査	筑波大学	教授	博士（教育学）	芳鐘冬樹
副 査	筑波大学	准教授	博士（工学）	宇陀則彦
副 査	千葉大学	教授	文学修士	竹内比呂也

論 文 の 要 旨

近年、インターネットを介した世界規模での研究データ共有や研究データ公開が進んでいる。研究データ公開と、学術論文をインターネット上で制限なく公開するオープンアクセス(OA)を含むオープンサイエンス(OS)は、学術情報政策、研究の効率化、研究不正への対応等、今日の学術情報流通において重要な案件になっている。本論文は日本の研究者によるデータ公開の実践状況および研究者の認識と課題を明らかにすることを研究目的としている。そのために、科学技術専門家ネットワーク 1,983 名への Web 質問紙調査、および 22 分野の高インパクトファクター学術雑誌を対象としたデータ公開ポリシー調査を実施し、分野別にデータ公開要求の状況と日本の研究者の実践状況を比較分析している。

本論文は5章から構成されている。

第1章では研究背景として研究データ公開の推進にかかわる政策、学術雑誌のポリシーの制定状況等を記述している。それを踏まえ日本の研究者によるデータ公開状況およびデータ公開に対する認識と課題を明らかにすることを研究目的として示している。研究課題として(1)日本の研究者はどの程度データを公開しているのか、(2)日本の研究者によるデータ公開の内的要因は何か、(3)データ公開の外的要因はどのような状況にあるのか、この外的要因と日本の研究者によるデータ公開状況に乖離はあるのか、の3点を設定している。研究方法として日本の研究者を対象とした質問紙調査と学術雑誌のデータ公開ポリシー調査を行うことを記している。

第2章では文献レビューとしてデータと研究の関わり、データ公開の意義と効果、データ公開を推進する取り組み、データ公開に関する実態調査、学術雑誌のデータ公開ポリシーに関する先行研究および政策文書について述べている。

第3章では日本の研究者によるデータ公開の状況と内的要因を明らかにするために、2016年11月から12月にかけて Web を通じた質問紙調査を実施し、その結果と考察を記している。文部科学省科学技術・学術政策研究所 (NISTEP) が運営する科学技術専門家ネットワークに登録されている大学、企業、公的機関・団体に所属する研究者 1,983 名を対象に、35 問の質問によってデータ公開経験、公開データの利用経験、論文を OA とした経験およびデータ公開の動機や障壁などについて尋ね、1,398 名 (70.5%) から有効回答を得ている。この結果を Clarivate Analytics 社の研究動向・統計データベース Essential Science Indicators (ESI) に基づき全 22 分野に分け、OA の経験、所属機関、年齢、データ公開経験の有無等を比較項目として分析している。

日本の研究者のデータ公開率は、51.0%であるとの結果を得ている。所属機関別にみたデータ公開率は大学 (56.9%)、公的機関・団体 (55.6%)、企業 (32.4%) であったとしている。一方、論文を OA とした経験をもつ回答者の割合は 70.9%であったとしている。

データを公開する理由の上位は「研究成果の認知度向上」(58.5%)、「雑誌のポリシー」(43.8%) であった。以下、「科学研究や成果実装の推進」(26.1%)、「他の研究者からのリクエスト」(25.4%) 「所属機関のポリシー」(23.7%) 「オープンデータ/OA への貢献」(10.5%) 「分野・コミュニティの規範」(9.5%) 「業績になる場合がある」(9.3%) 「助成機関のポリシー」(6.3%) であったとしている。

非公開理由は「雑誌のポリシーでない」(26.4%) 「時間が必要」(25.9%) 「所属機関にポリシーがない」(22.4%) 「ニーズがないと思うから」(22.1%) 「業績にならないから」(21.7%) の結果が得られている。その他、データの公開と OA の意思、データ公開する場合の資源状況、データ公開する場合の懸念等を調査しており、これをもとに考察を行っている。

第4章では学術雑誌のデータ公開ポリシーについて調査を行っている。ESI の雑誌リストを用いて、22 分野からインパクトファクターの高い原著論文誌を 10 誌ずつ抽出して調査対象とし、合計 220 誌の投稿規定などを調査し、リポジトリにデータを公開して論文に識別子を記すポリシー (Repository Policy, 以下「RP」) と、補足資料にデータを掲載するポリシー (Supplemental Materials' Policy, 以下「SP」) を、要求の強度すなわち「必須 (require)」、「推奨 (recommend)」、「受諾 (accept)」、「なし (no mention)」に従って 4 段階に分類し、分析を行っている。また雑誌のダイレクトリである Ulrichsweb.com などを用いて各学術雑誌の特徴を調査し、その上で分野ごとの RP の掲載率と調査回答者によるデータ公開率を比較している。その結果、調査対象である 220 誌のうち何らかのデータ公開ポリシーを掲載していた雑誌は 204 誌 (92.7%) であったとしている。また RP の掲載率は 59.5%、SP の掲載率は 89.5% であったとしている。分野別では生物学と医学の関連領域は RP の掲載率や強度が高い結果を示している。

調査対象誌の特徴として、(1) 倫理や利益相反 (Conflict of Interest, COI) の開示に関する記述の有無、(2) 出版者 (学協会と商業出版社)、(3) OA ステータスを調査している。(1) 倫理や COI の開示に関する記述がある雑誌は、ポリシーの掲載率や要求の強度が高いこと、(2) 出版者は商業出版社の方が掲載率も要求の強度も高かったことから、学術雑誌によるデータ公開ポリシーは研究不正に関連がある可能性を指摘している。(3) OA ステータスについて、OA 誌の方がデータ公開を強く要求しているという結果は得られていないとしている。

また、22 分野のデータ公開要求の状況と、第3章で明らかにした日本の研究者によるデータ公開状況を比較して、両者に相関があることを示している。その上で、データ公開要求に対してデータ公開率が高い分野を「自発的にデータ公開を行っている分野」、データ公開率が低い分野を「データ公開が困難な分野」として、他分野との比較からデータ公開において障壁が低い点や高い点をそれぞれ示している。

第5章では3章4章での結果と考察を基に研究課題への回答を示し、総合的な考察と結論を述べている。そこではOSにかかわる出版者、資金提供機関、研究機関等のステークホルダーが対策を講じる観点からデータ公開に関する課題が整理され、データ公開が緒に就いた段階であると指摘している。また、データ公開が適さない場合を示し、一律にデータ公開を求めることがないように留意すべきであると指摘している。

審 査 の 要 旨

【批評】

本研究は、公的資金による研究の成果である論文とデータを公開して市民が自由にアクセスできるようにするためのオープンサイエンス(OS)に関わる日本の研究者によるデータ公開の実践状況と認識を明らかにすることを研究目的としている。学術雑誌価格の高騰による学術論文のアクセス制限をインターネットで論文を利用できるようにするオープンアクセス(OA)問題と並び、学術論文の根拠となる研究データ公開を進めようとするOSの現状とOSに対する研究者の意識を明らかにすることは今日の学術情報流通において重要な課題である。この課題に取り組んだ本研究はオープンサイエンスの進捗と課題について貴重な成果をもたらし、重要な示唆を与えるものである。

研究課題(1)では日本の研究者のデータ公開の現状を、研究課題(2)では日本の研究者によるデータ公開の内的要因としてデータ公開の理由、障壁等を、研究課題(3)ではデータ公開の外的要因として学術雑誌のデータ公開ポリシーを調査している。これは妥当な設定である。データ公開の実態について国内において小規模な調査はあるが、本研究では文部科学省科学技術・学術政策研究所(NISTEP)の協力を得て調査を行い、1,398名から回答を得て、主題、属性、OAへの意識など多面的な結果を導出し考察を行っている。またデータ公開に強い影響を与えるとされるデータ公開に関わる学術雑誌のポリシーを明らかにしている。22分野からインパクトファクターの高い原著論文誌を10誌ずつ抽出して調査対象とし、合計220誌の投稿規定などを調査し、リポジトリにデータを公開して論文に識別子を記すポリシーと、補足資料にデータを掲載するポリシーの要求の強度を明らかにし、分野別のデータ公開要求の状況をマクロに捉えようとしている。その結果を基に結論を導いている。これだけの規模で取り組んでいる先行研究はなく、学術情報政策や研究データ管理の実践に重要な内容を示していると評価できる。

第1章では研究背景、研究目的、研究課題、研究方法、研究の学術的な意義と社会的な意義について述べている。研究データ公開にかかわる政策、学術雑誌のポリシーの制定状況等を示している。それらを踏まえたここでの記述は丁寧に行われており、適切である。

第2章では、研究課題に関係する文献をレビューしている。まず、研究の背景となるデータと研究の関わり、データ公開の意義と効果、データを推進する取り組みについてレビューしている。続いてデータ公開に関する実態調査、データ公開の動機と障壁、雑誌のデータ公開ポリシーにかかわる文献をレビューした上で先行研究では未解明の点を記述している。第3章と第4章の調査方法はこれらの先行研究を批判的に検討した上で設計されており、本論文に関係する文献レビューとしては十分なものと評価できる。

第3章では日本の研究者におけるデータ公開の状況と内的要因を明らかにするためにNISTEPの協力

を得て調査を行っている。主題、属性、論文を OA にした経験、データ公開の経験と意思、公開データの利用経験データ公開の動機と障壁など多面的な結果を導出し考察している。1,398 名に及ぶ調査は丁寧に行われており、それぞれの項目について興味深い結果が得られている。また、回答者をデータ公開経験の有無やデータ公開意思の有無により 2 群に分けて、両者の比較からそれぞれの特徴を抽出している。大学のみならず公的機関・団体や、データ公開の事例が増えつつある企業の研究者にも調査を行い、比較分析によってそれぞれの特徴を示した点も研究のオリジナルな成果であるといえる。ただ回答者を 22 分野に分けて結果を示しているが、分野によって回答人数の差が大きく、結果にバイアスがやや生じている点は否めない。

第 4 章では学術雑誌のデータ公開ポリシーの状況を分野別に明らかにするために、高インパクトファクター原著論文誌を 22 分野から各 10 誌、計 220 誌を選び、その投稿規定などを調査し、データ公開にかかわるポリシーの強度を調べ、分野ごとの特徴を明らかにしている。対象誌の選定は先行研究を参考とした予備調査を経て行われ、研究倫理、利益相反、出版者、OA ステータスについて分析を行い、考察を行っている。これまでデータ公開ポリシーと実際の公開率との関連はほとんど調査されておらず、研究の独自性として評価できる。また、日本の研究者による分野別のデータ公開状況を相対的に明らかにするために、第 3 章の結果と比較し考察を行っている。第 3 章での調査対象者が認識している雑誌のポリシーの要求状況（内的要因）と第 4 章で対象とした学術雑誌のポリシーの要求状況（外的要因）の関係についてはより詳細な説明が望まれる。

第 5 章では、第 3 章と第 4 章の結果考察を基に総合的な考察と結論を導いている。第 5 章の総合的な考察や結論は学術情報政策の観点が中心となっている観があるが、全体としてオリジナルな貢献をし、研究データ公開の現状と将来において重要な示唆を与えるものであると考えられる。特に日本の研究者によるデータの公開状況と課題を明らかにし、国外の状況と相対的に比較できるようにしたことは、研究の意義として高く評価できる。変化の激しいテーマであり、継続的な調査や人文学分野を含めた調査を行うことが今後に残された課題であるといえる。

以上を総合的に判断すると、本論文は図書館情報学の学位論文として十分な内容を有すると認められる。

【学力の確認結果】

令和元年 8 月 5 日、図書館情報メディア研究科学位論文審査委員会において、審査委員全員出席のもと、本論文について著者に説明を求めた後、関連事項について質疑応答を行った。引き続き、「図書館情報メディア研究科博士後期課程（論文博士）の学位論文審査に関する内規」第 23 項第 3 号に基づく学力の確認を行い、審議の結果、審査委員全員一致で合格と判定された。

【結論】

よって、本学位論文の著者は博士（図書館情報学）の学位を受けるに十分な資格を有するものと認める。